

EP 99  
6847



PCT/EP 99/06847

EPO - Munich  
8

24. Sep. 1999

EU

# ÖSTERREICHISCHES PATENTAMT

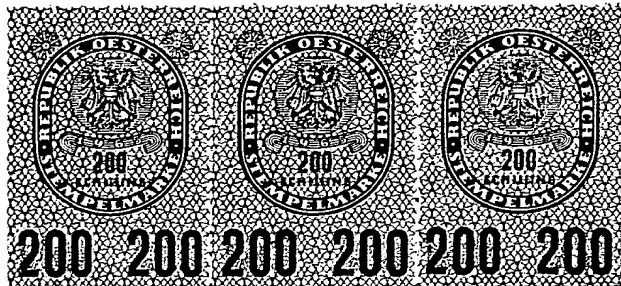
A-1014 WIEN, KOHLMARKT 8 - 10

REC'D 12 NOV 1999

WIPO

PCT

Aktenzeichen A 1625/98



Das Österreichische Patentamt bestätigt, dass

**die Firma Siemens Aktiengesellschaft Österreich  
in A-1210 Wien, Siemensstraße 92,**

**am 30. September 1998 eine Patentanmeldung betreffend**

**"Verfahren zur Zeichentrennung bei Texterkennungsaufgaben",**

überreicht hat und dass die beigeheftete Beschreibung samt Zeichnung mit der ursprünglichen, zugleich mit dieser Patentanmeldung überreichten Beschreibung samt Zeichnung übereinstimmt.

**PRIORITY  
DOCUMENT**

SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

Österreichisches Patentamt  
Wien, am 25. August 1999

Der Präsident:

i. A.



Kanzleirat FUHLINGER  
Fachoberinspektor



ÖSTERREICHISCHES PATENTAMT  
Verwaltungsstellen-Direktion

.....*20*..... S ..... €

Kanzleigebühr bezahlt.

*Balaban*

A 1625 / 98-1

Urtext

Siemens Aktiengesellschaft Österreich

Pm

Wien, 30. 09. 1998

GR 98 P 7018 AT

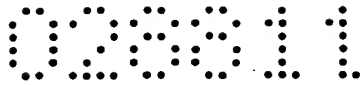
### Verfahren zur Zeichentrennung bei Texterkennungsaufgaben

Die Erfindung betrifft ein Verfahren zur Zeichentrennung bei Texterkennungsaufgaben.

Bei der automatischen Erkennung von Texten, d.h. bei der Umwandlung der grafischen Information eines Dokumentes in Textzeichen, die mittels elektronischen Textverarbeitungsprogrammen weiterverarbeitet werden können, ist eine wesentliche Voraussetzung für einen erfolgreichen Erkennungsvorgang die genaue Bestimmung der Lage und der Größe der einzelnen Zeichen. Diese Bestimmung ist bei Vorlagen mit schlechtem Schriftbild oder Schriftarten mit sehr engem Zeichenabstand unter anderem dadurch problematisch, daß die Zeichen mit einander verbunden sind "zusammenwachsen" und damit durch herkömmliche Methoden wie die einfache Konturverfolgung nicht mehr getrennt werden können.

Der Erfindung liegt daher die Aufgabe zugrunde, ein verbessertes Verfahren zur Trennung miteinander verbundener Zeichen anzugeben.

Dies geschieht erfindungsgemäß mit einem Verfahren der eingangs genannten Art, bei dem zu den untersuchten Extraktionsobjekten mittels Weißdellenanalyse und Winkelanalyse mögliche Schnittpunkte ermittelt werden, bei dem aus den Schnittpunkten und entsprechenden Gegenpunkten plausible Trennlinien ermittelt werden und bei dem die solcherart getrennten Objekte Klassifikationsverfahren unterzogen werden und auf der Grundlage der Ergebnisse die endgültige Trennung erfolgt.



Vorteilhaft ist eine Ausgestaltung des Verfahrens in der Weise, daß bei mehr als drei möglichen Schnittpunkten, ein erster Schnitt durch den vom linken Zeichenanfang gezählten vierten Schnittpunkt erfolgt. Dies deswegen, weil kein übliches Textzeichen der lateinischen Schrift mehr als drei Weißdellen aufweist.

Günstig ist es ferner, wenn nach einem ersten Schnitt mit einem ersten möglichen Schnittpunkt und einem darauffolgenden erfolglosen Klassifikationsversuch als Basis für einen weiteren Trennversuch der zum ersten möglichen Schnittpunkt nächstliegende linke Nachbarschnittpunkt vorgesehen wird.

Die Erfindung wird anhand von Figuren näher erläutert.  
Es zeigen beispielhaft:

Fig.1 eine Darstellung zur Weißdellenanalyse eines Bildes,  
Fig.2 eine Darstellung zur eigentlichen Zeichentrennung.

Der Ablauf des erfindungsgemäßen Verfahrens ist wie folgt:

Das Verfahren wird im Erkennungsvorgang nach der Bestimmung der Lage der Zeile gestartet. Bei der Ermittlung des Umfanges eines Zeichens oder mehrerer verbundener Zeichen durch Konturverfolgung wird bereits eine Weißdellenanalyse durchgeführt. Nach dem Vorliegen der vollständigen Kontur erfolgt eine Winkelanalyse.

Mittels Weißdellenanalyse und Winkelanalyse werden mögliche Schnittpunkte ermittelt, die in Verbindung mit Gegenpunkten mögliche Trennlinien liefern.

Die Schnittpunkte werden hinsichtlich ihrer Plausibilität untersucht. Dabei wird ermittelt, welche Zeichenfolgen die vorliegende Weißdellenkombination beinhalten. So sind beispielsweise in der Buchstabenfolge WV folgende

Weißdellen enthalten OBEN-UNTEN-OBEN-UNTEN-OBEN. Wobei OBEN (UNTEN) eine nach oben (unten) offene Weißdelle kennzeichnet. Aus der Kenntnis der Buchstaben heraus wird nun die erste Trennung durch den Schnittpunkt der vierten Weißdelle erfolgen.

Darauf wird ermittelt, inwieweit die Trennung des Objekts entlang der auf plausiblen Schnittpunkten beruhenden Trennlinien zu plausiblen Klassifikationsergebnissen führt. Mit anderen Worten, die getrennten Zeichen oder Zeichenteile werden einem Erkennungsvorgang z.B. mittels neuronalem Netz unterworfen und wenn dieser Vorgang zu einem zufriedenstellenden Ergebnis - einem mit hoher Sicherheit erkannten Zeichen - führt, dann wird die Trennung akzeptiert. Andernfalls wird die Trennung entlang von anderen Trennlinien solange wiederholt, bis ein zufriedenstellendes Ergebnis vorliegt.

Neuronale Netze sind mathematische Modelle, welche dem Aufbau des menschlichen Gehirns nachempfunden sind. Sie bestehen aus Neuronen, das sind im wesentlichen Summierelemente mit gewichteten Eingängen und einem nichtlinearen Verstärkeranteil, die zu einem parallelen Netzwerk mit typisch zwei Ebenen zusammengefaßt werden. Eine ausführliche Beschreibung des beim Ausführungsbeispiel eingesetzten "Feedforward Neural Networks" findet sich beispielsweise in "Layered Neural Nets for Pattern Recognition", B. Widrow, R. G. Winter, R. A. Baxter; IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. 36. No. 7. July 88.

Die Mustererkennung mittels neuronalem Netz erfolgt nach dem in "A rotation, scaling, and translation invariant pattern classification system", C. Yüceer, K Oflazer; Pattern Recognition, Vol. 26, No5, pp687-710, 1993. beschriebenen Verfahren.

Die Weißdellenanalyse wird anhand der Figur 1 näher beschrieben. Die Figur zeigt die beiden miteinander verbundenen Buchstaben r und f die eine Weißdelle W aufweisen. Unter Weißdelle W wird dabei ein von drei Seiten begrenzter weißer Zwischenraum verstanden, der eine gewisse Tiefe aufweist und dessen offene Seite nach oben oder unten gerichtet ist. Ermittelt wird diese Weißdelle W bei der Verfolgung der Kontur des (zusammengewachsenen) Zeichens wenn die Konturlinie C zwei vorgegebene Schwellwerte SW in der richtigen Reihenfolge in beiden Richtungen überschreitet. Liegt wie in dem Beispiel eine nach unten offene Weißdelle W vor, dann wird der höchste Punkt der Konturlinie C als möglicher Schnittpunkt S definiert, bei einer nach oben offenen Weißdelle ist dies der tiefste Punkt.

Der Ablauf der daraufhin erfolgenden Winkelanalyse ist wie folgt:

Aus jeweils drei Punkten der Konturlinie  $C[i]$  werden zwei Vektoren ermittelt, für die gilt:

$$\vec{A} = C[i]C[i-5] \text{ und } \vec{B} = C[i]C[i+5]$$

Der Winkel zwischen den beiden Vektoren wird berechnet. Ist dieser linksläufig, mit einem Betrag kleiner als  $80^\circ$  und einer entweder nach oben oder nach unten weisenden Spitze ( $C[i]$ ), dann wird der Winkel in eine Liste eingetragen.

Ist diese Bedingung für mehrere nebeneinander liegende Vektorpaare erfüllt, dann wird nur der Winkel mit dem geringsten Betrag weiterverfolgt.

Die in der Liste eingetragenen Winkel werden nun daraufhin untersucht, ob auf der gegenüberliegenden Seite der

Konturlinie ein Winkel mit entgegengesetzter Orientierung der Spitze vorhanden ist. Ist dies der Fall, dann wird das daraufhin gebildete Winkelpaar als Position eines möglichen Schnittpunktes gespeichert.

Im Folgenden der Ablauf bei der Bestimmung des Winkels zwischen zwei Vektoren, die durch 3 Punkte aus der Konturlinie ( $C_1:x_1/y_1$ ,  $C_6:x_2/y_2$ ,  $C_{11}:mx/my$ ) definiert sind. Daraus werden die x und y Komponenten der beiden Vektoren ermittelt.

$$Ax = x_1 - mx ; Ay = y_1 - my ; Bx = x_2 - mx ; By = y_2 - my ;$$

Der Winkel zwischen den Vektoren A und B wird wie folgt berechnet: Zuerst wird der Winkel von A zur x-Achse und dann der Winkel B zur x-Achse ermittelt.

$$Winkel = \arccos\left(\frac{\bar{A}x}{\sqrt{(\bar{A}x)^2 + (\bar{A}y)^2}}\right)$$

$$Winkel(inGrad) = \frac{Winkel(inRad) * 180}{\pi}$$

Winkel = 360-winkelB+winkelA (ist Winkel größer 360°, dann wird der Winkel um 360° korrigiert)

Die Bestimmung der Winkelspitzenrichtung beruht auf der Überlegung, daß bei einer nach unten gerichteten Spitze die Y-Koordinaten der Punkte  $C_1$  und  $C_6$  kleiner als die Y-Koordinate von  $C_{11}$  sind.

Bei einer nach oben gerichteten Spitze müssen hingegen die Y-Koordinaten der Punkte  $C_1$  und  $C_6$  größer als die Y-Koordinate von  $C_{11}$  sein.

Die Eigenheiten gedruckter Texte und der Einfluß der begrenzten Bildauflösung bringen es mit sich, daß im Bereich eines Knicks der Kontur eines Zeichens die in der beschriebenen Weise ermittelten Winkel zwischen 2 Vektoren abhängig vom Betrachtungsraum zuerst zunehmend kleiner werden und danach wieder kontinuierlich zunehmen. Für die weitere Auswertung wird daher nur der jeweils minimale Winkel eines derartigen Bereiches verwendet.

Zur Festlegung einer möglichen Trennlinie muß nun zu jedem möglichen Schnittpunkt  $C(Nr)$  ein entsprechender Gegenpunkt auf dem gegenüberliegenden Zweig der Konturlinie  $C(i); i=(0, \dots, \text{contourNr})$  ermittelt werden.

Dazu wird eine Gerade durch zwei auf der Konturlinie dem möglichen Schnittpunkt  $C(Nr)$  benachbarte Punkte  $C(Nr-1)$  und  $C(Nr+1)$  gelegt, und zu dieser Geraden die Normale ermittelt. Die zu dem Schnittpunkt dieser Normalen mit dem gegenüberliegenden Zweig der Konturlinie benachbarten Punkte werden hinsichtlich ihres Abstandswertes zum möglichen Schnittpunkt und der Normalen untersucht und der Konturpunkt mit dem minimalen Abstandswert als Gegenpunkt  $C(g)$  und damit als zweiter Punkt der möglichen Trennlinie definiert.

Die mathematische Definition dieses Vorganges lautet:

$$nx = C(Nr+1)x - C(Nr-1)x$$

$$ny = C(Nr+1)y - C(Nr-1)y$$

$$\text{Abstand} = \sqrt{(C(Nr)x - C(i)x)^2 + (C(Nr)y - C(i)y)^2}$$

$$\text{Abstand zu } g2 = \text{abs} \left( \frac{nx * (C(i)x - C(Nr)x) + ny * (C(i)y - C(Nr)y)}{\sqrt{nx^2 + ny^2}} \right)$$



Abstandswert = Abstand + Abstand zu g2;

$C(g) = C(i) \mid \text{Abstandswert}(C(g), C(Nr)) = \min$

Die eigentliche Trennung wird anhand der Fig. 2 erläutert: Basis der Trennung ist die Konturlinie der extrahierten Zeichen. In einem 1. Schritt wird ein Trennlinienpuffer mit 0 initialisiert, dies entspricht einer senkrechten Linie am linken Rand, danach wird der am weitesten rechts liegende Punkt der Konturlinie 1 zwischen 0 und dem der Trennung zugrundeliegenden Schnittpunkt (das X-Wert-Maximum) ermittelt. Ebenso werden der am weitesten rechts liegende Punkt des Zweiges (das x-Wert Maximum) der Konturlinie vom Gegenpunkt bis zum Ende der Kontur 2 und der Trennlinie 3 ermittelt.

Die gesammelten maximalen x-Werte stellen also den äußersten rechten Rand des zur Klassifizierung herangezogenen Zeichens dar.

Patentansprüche

- 1) Verfahren zur Zeichentrennung bei Texterkennungsaufgaben, **dadurch gekennzeichnet**, daß zu den untersuchten Extraktionsobjekten mittels Weißdellenanalyse und Winkelanalyse mögliche Schnittpunkte ermittelt werden, daß aus den Schnittpunkten und entsprechenden Gegenpunkten plausible Trennlinien ermittelt werden und daß die solcherart getrennten Objekte Klassifikationsverfahren unterzogen werden und auf der Grundlage der Ergebnisse die endgültige Trennung erfolgt.
- 2) Verfahren nach Anspruch 1, **dadurch gekennzeichnet**, daß bei mehr als drei möglichen Schnittpunkten, ein erster Schnitt durch den vom linken Zeichenanfang gezählten vierten Schnittpunkt erfolgt.
- 3) Verfahren nach Anspruch 1 oder 2, **dadurch gekennzeichnet**, daß nach einem ersten Schnitt mit einem ersten möglichen Schnittpunkt und einem darauffolgenden erfolglosen Klassifikationsversuch als Basis für einen weiteren Trennversuch der zum ersten möglichen Schnittpunkt nächstliegende linke Nachbarschnittpunkt vorgesehen wird.

Zusammenfassung:

Es wird ein Verfahren zur Zeichentrennung bei Texterkennungsaufgaben angegeben, bei dem zu den untersuchten Extraktionsobjekten mittels Weißdellenanalyse und Winkelanalyse mögliche Schnittpunkte ermittelt werden, daß aus den Schnittpunkten und entsprechenden Gegenpunkten plausible Trennlinien ermittelt werden und daß die solcherart getrennten Objekte Klassifikationsverfahren unterzogen werden und auf der Grundlage der Ergebnisse die endgültige Trennung erfolgt.

*Fig. 1*

Siemens Aktiengesellschaft Österreich  
durch



Handlungsvollmacht

Zl. 2275 /Präs. 96

A1625/98-1:08811

Urtext

GR 98P7018 AT

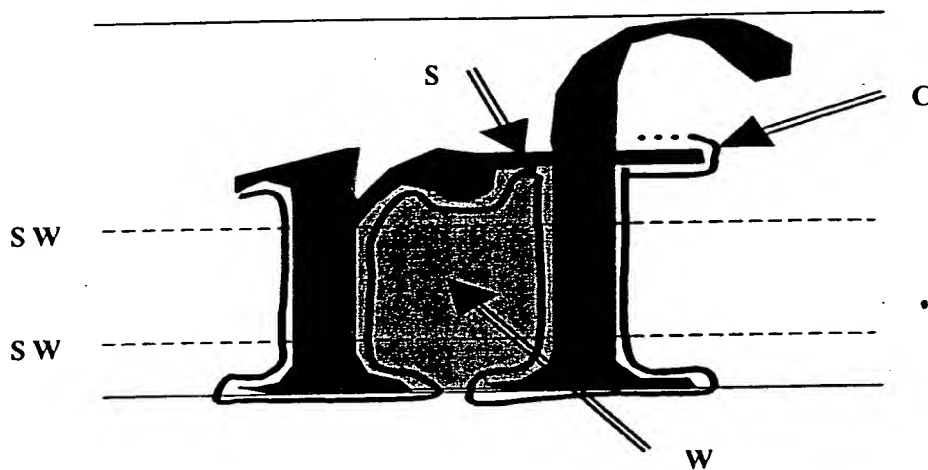


Fig. 1

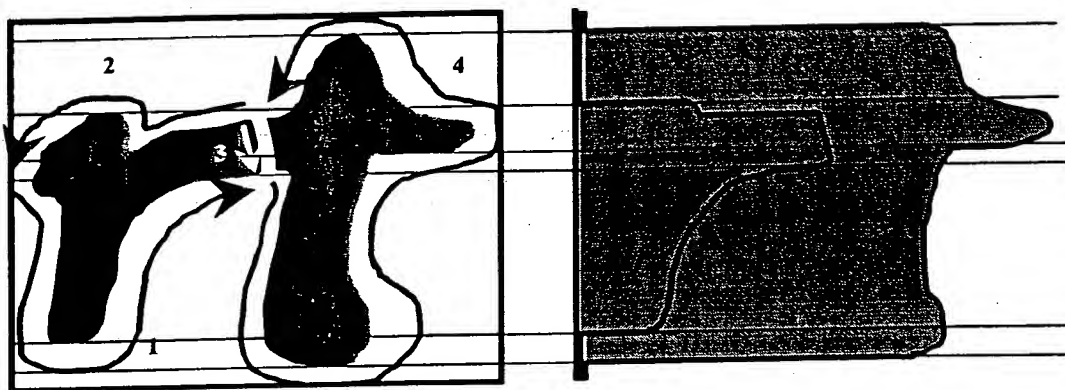


Fig. 2

Siemens Aktiengesellschaft Österreich  
durch

*iv KLC*

Handlungsvollmacht  
Zl. 2275 /Präs. 96